

No Regrets agent

**for ANAC SCML
2021**

TheSopranos78

Moshe Binieli
Tal Tikochinski
Ran Ezra

mmoshikoo@gmail.com
taltiko26@gmail.com
Ran421@gmail.com

No Regrets agent

The scenario

The agent works in a manufacturing plant, and does in each session one of the following two things:

1. Works in sales and needs to sell as many of the products that the company produces (the company produces one product in varying quantities) at the highest price it can.
2. Works in procurement and needs to purchase materials that will be enough for the company to produce the number of products they have committed to, of course at the cheapest price possible.

It is required to create a strategy that will suit the 2 scenarios, because the agent switches between the roles randomly.

What does the agent do?

Monitors market prices regularly and operates in a greedy way, **but** not entirely greedy.

Greedy in what way?

- Trying as much as possible to avoid a situation where I "regret":
 - An offer that was offered to me and refused it (because they only got worse later)
 - A counter-offer that was too low that I made but could have asked for and received more
- Will always ask for the best when I am the one offering (the quantity I need at the best price as for the current time)
- Will immediately take a relatively good offer, without thinking about improving the offer in the future (taking any quantity equal to or less than the quantity I need if I think the price is good now)

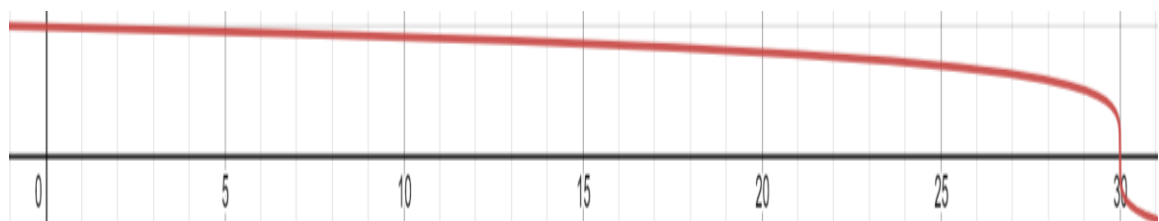
Not entirely greedy in what way?

- Willing to make small compromises along the way and big compromises just right towards the end. This means that if I sell, I know what the maximum possible price is
- Willing to get off it just a little bit for most of the negotiations and a lot towards its end. If I buy, I know the minimum and am willing to go up a little at the beginning and a lot at the end

How the agent does that?

We define a function that will determine how much I am willing to move from the best price at each stage in the negotiation, depending on the stage I am in, and while knowing how many steps there are at the maximum (neg steps). The function is a fading exponent when the estimator (a locally calibrated hyperparameter) is less than 1 which causes it to drop slowly most of the time and fast at the end (the lower the function the way I move farther from the best price).

$$\left(\frac{\text{neg steps} - \text{current step}}{\text{neg steps}} \right)^{\text{param } e < 1}$$



I know the required quantities because I have communication with the requirements that I need for me (contracts signed with suppliers and customers) and I also know how much and maximum the agent who negotiates with me can sell/buy. I will of course not ask a dangerous person to sell me more or less than he can because it will be probably a failure.

Why it works and when will it not be good?

First of all, and this is also the most important - we see the results of our agent in the tournament and that proves its performance.

When the agent may not be good? If some agent whose negotiating with our agent for long periods while utilizing the maximum time, he can observe our behavior and take advantage of it; meaning this agent will notice that towards the last steps our agent can systematically down-price the offer so this agent can delay the final offer and maximize his profit.

Our assessment is that no agent will reach this level of experience when negotiating with our agent because in order to notice a change he has to reach a very advanced stage many times (parallel to searching for a solution that is very deep in the search space, learning, in this case, requires a lot of experience).